DOCUMENT RESUME

ED 276 761 TM 860 712

AUTHOR Tracz, Susan M.; And Others
TITLE Tests of Dependence in Meta-Analysis Using Multiple Linear Regression.
PUB DATE Apr 86
NOTE 11p.; Paper presented at the Annual Meeting of the American Educational Research Association (67th, San Francisco, CA, April 16-20, 1986).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Effect Size; *Mathematical Models; *Meta Analysis; *Multiple Regression Analysis; Research Design; Research Methodology; Statistical Significance; Statistical Studies
IDENTIFIERS Data Dependence; Multiple Linear Regression

ABSTRACT

The purpose of this paper is to demonstrate how multiple linear regression provides a viable statistical methodology for dealing with meta-analysis in general, and specifically with the issues of nonindependence and design complexity, such as multiple treatments. Since the F-test and t-test are special cases of the general linear model, approaching meta-analysis from a regression perspective allows the researcher much more flexibility in the questions that can be asked, while insuring that the appropriate error terms and degrees of freedom are used. A number of full and restricted mathematical models are illustrated. A test of effect sizes over and above the dependency of data is presented. (GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED276761

Tests of Dependence in Meta-Analysis Using Multiple Linear Regression

Susan M. Tracz
California State University, Fresno

Isadore Newman
University of Akron

Keith McNeil
Dallas Unified School District

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

S. M. Tracz

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

A paper presented at the annual meeting of the American Educational Research Association, April, 1986

TM 860 712

BEST COPY AVAILABLE

Meta-analysis, per se, is a combinatory methodology developed by Glass (1976) to synthesize the independent, statistical results of primary research studies. It follows in the history of combinatory methodology for probabilities (Rosenthal, 1978) and Cohen's effect size (1977). Traditionally, meta-analyses are reported in terms of the mean effect size, $\overline{ES}$, where effect size is defined as the difference in means between experimental and control groups divided by a pooled standard error,

$$ES = \frac{\overline{X}_E - \overline{X}_C}{S_x} \qquad (1)$$

This type of result, of course, is purely descriptive in nature.

Meta-analysis is becoming an increasingly popular statistical technique. A plethora of meta-analyses appear in the literature, and some federal agencies now require that a meta-analysis be performed before a project is considered for a grant. However, meta-analysis is still an evolving statistical technique. New advances are the inclusion of multivariate studies (Busk & Marascuilo, 1985) and the use of regression techniques (Hedges & Olkin, 1985).

Two important issues have arisen in the application of meta-analysis. First, the issue of nonindependence of data has been raised but never dealt with comprehensively. If the assumption of independence were met, only one effect per study would be combined (Kraemer, 1983). However, even a cursory review of meta-analyses would reveal that in practice this is rarely done. Some authors have addressed nonindependence. Strube (1985) suggests procedures for adjusting for nonindependence when combining significance levels, but the necessary data is seldom reported in journal articles. Tracz and Elmore (1985) performed a Monte Carlo simulation which indicated that violating the assumption of independence did not affect the results when combining correlation coefficients. However, these solutions are either not practical or not general enough to settle the issue of nonindependence.

Second, not all studies have an experimental versus control designs. More complicated designs such as those reporting more than one treatment are much more common.

The purpose of this paper is to demonstrate how multiple linear regression provides a viable statistical methodology for dealing with meta-analysis in general and with the issues of nonindependence and design complexity specifically. Since the F and t-tests are special cases of the general linear model, approaching meta-analysis from a regression perspective allows the researcher much more flexibility in the questions that can be asked while insuring that the appropriate error terms and degrees of freedom are used.

Instead of merely reporting the mean for a vector of effect sizes, ES (1), a test for differences between the experimental and control groups or for the direction of the effect sizes can be performed by running a full model:

$$ES = b_0U + b_1X_1 + E \qquad (2)$$

where ES = the absolute values of the effect sizes as defined in equation 1,

b_0, b_1 = the least squares regression coefficients,

U = the unit vector,

X_1 = a vector coded 1 if the effect size is positive, 0 otherwise, and

E = the error term.

The restricted model then would be:

$$ES = b_0U + E \qquad (3)$$

Finally, the F test can be calculated as follows:

$$F = \frac{(R_F^2 - R_R^2)/\ df_1}{(1 - R_F^2)/\ df_2} \qquad (4)$$

where df_1 = the number of weights in the full model minus the number of weights in the restricted model,

df_2 = N – the number of weights in the full model,

N = the total number of effect sizes.

This analysis parallels the independent t-test performed with multiple linear regression (Cohen & Cohen, 1975; McNeil, Kelly & McNeil, 1975; Pedhazur, 1982). Here the full model is:

$$Y = b_0U + b_1X_1 + E \qquad (5)$$

where Y = the dependent variable,

b_0, b_1 = the least squares regression coefficients,

U = the unit vector,

X_1 = a vector coded 1 if the subject is in experimental group, 0 for the control group, and

E = the error term.

The restricted model then would be:

$$Y = b_0 U + E \qquad (6)$$

and F-test (4) is the same as above.

Another way of stating this full model (5) is given below:

$$Y = b_1 X_1 + b_2 X_2 + E \qquad (7)$$

where Y = the dependent variable,

b_1, b_2 = the least squares regression coefficients,

X_1 = a vector coded 1 if the subject is in the experimental group, 0 otherwise,

X_2 = a vector coded 1 if the subject is in the control group, 0 otherwise, and

E = the error term.

In this case the solution for b_1 is the mean of the experimental group and the solution for b_2 is the mean of the control group.

If a series of N-1 subject or person vectors is added to the full model (5), where the vector X_1 is coded 1 for the pretest and 0 for the posttest, and to the restricted model, the test becomes a dependent t-test (Kelly, McNeil, Newman, 1971; McNeil, Kelly, & McNeil, 1975). If a series of N-1 subject or person vectors is added to the full model, where the time factor has three or more levels, and to the restricted model, the test becomes a repeated measures design (Fanning & Newman, 1974; Pedhazur, 1982). This is because the dependent t-test is the simplest case of a repeated measures design. These full models can be compared to models restricted to test for group differences or subject differences.

If a series of dummy coded vectors for studies instead of subjects is included in the full model, the issue of dependence can be addressed in a similar manner. This parallels work done on dependence of data in chi-square analysis (Newman, Williams, & Bobner, 1982). The following full model could be used in the situation where one experimental and one control group were being compared in three studies:

$$ES = b_oU + b_1X_1 + b_2S_1 + b_3S_2 + E \qquad (8)$$

where ES = the absolute values of the effect sizes,

b_o, b_1, b_2, b_3, = the least squares regression coefficients,

U = the unit vector,

X_1 = a vector coded 1 if the effect size is positive, 0 otherwise,

S1, S2 = the dummy coded vectors for the three studies, and

E = the error term.

A restricted model, setting the weights for the treatment vector equal to zero, would then be:

$$ES = b_oU + b_2S_1 + b_3S_2 + E \qquad (9)$$

An F test of the difference between experimental and control treatments controlling for study dependence could then be performed using formula 4. This test controls for dependence by using the appropriate error term and degrees of freedom.

A test of study dependence could be run by including only dummy coded vectors for studies in the full model. For example, regardless of the number of treatments, if the data from three studies was included the full model would be:

$$ES = b_oU + b_1S_1 + b_2S_2 + E \qquad (10)$$

where ES = the effect sizes,

b_o, b_1, b_2 = the least squares regression coefficients,

U = the unit vector,

S_1, S_2 = the dummy coded vectors for the three studies, and

E = the error term

A restricted model (3) would set the weights for the study vectors equal to zero. An F test of study dependence could then be performed using formula 4. Again, this test uses the appropriate error term and degrees of freedom.

If the number of studies providing information to be combined became unwieldly then the following full model (Pedahuzer, 1982; Williams, 1983) could be run instead:

$$ES = b_0U + b_1X_1 + b_2M + E \qquad (11)$$

where ES = the absolute values of the effect sizes,

b_0, b_1, b_2 = the least squares regression coefficients,

U = the unit vector,

X_1 = a dummy coded vector for the sign of the effect size,

M = a vector containing means or sums of effect sizes within studies.

E = the error term

The restricted model would be:

$$ES = b_0U + b_2M + E \qquad (12)$$

The R^2 for the full models (8) and (11) and the R^2 for the restricted models (9) and (12) would be identical. However, when calculating the F test given in (4) the degrees of freedom for the full model (8) and the restricted model (9) should be used (Pedhazur, 1982). Alternately, the vector M could be coded with the sum of effect sizes within studies instead of the means (Williams, 1983). To illustrate this point, an example of coded data appears in Table 1.

Table 1
Data for Model 11

ES	X_1	S_1	S_2	M(means)	M(sums)
.2	1	1	0	.075	.3
.4	1	1	0	.075	.3
.1	0	1	0	.075	.3
.2	0	1	0	.075	.3
.5	1	0	1	.100	.4
.4	1	0	1	.100	.4
.3	0	0	1	.100	.4
.2	0	0	1	.100	.4
.4	1	0	0	.075	.3
.2	1	0	0	.075	.3
.2	0	0	0	.075	.3
.1	0	0	0	.075	.3

The problem of proportionality, however, must still be solved.

The model can be expanded to include the type of designs that more realistically appear in the literature, which are more sophisticated than simply having one experimental group and one control group. For example, if a control were being compared to two experimental treatments, the following full model would be appropriate:

$$ES = b_oU + b_1X_1 + b_2X_2 + B_3M + E \qquad (13)$$

where ES = the effect sizes,

b_o, b_1, b_2, b_3 = the least squares regression coefficients,

U = the unit vector,

X_1, X_2 = the dummy coded vectors for the two experimental conditions, and

M = a vector containing the means or sums within studies,

E = the error term.

In this case the restricted model would be:

$$Y = b_oU + b_3M + E \qquad (14)$$

and the F test is given in (4).

The model, of course, could be expanded to include terms coded to answer various research questions such as the relative importance of the quality of the study and the issue of trends of results over time. Topic specific terms could also be included. Additional, more sophisticated research questions could be addressed by including covariance, interaction, and curve-linear terms in the model (Fraar & Newman, 1978).

Developing and formalizing regression techniques for meta-analysis are important because researchers then would have more flexibility in analyzing effect sizes and would have a test of dependency of data. Finally, a test of effect sizes over and above the dependency of data is presented.

REFERENCES

Busk, P., & Marascuilo, L.A. (1985, March). Large-sample procedures for comparing and combining effect sizes within univariate and multivariate studies. Paper presented at the annual meeting of the American Statistical Association, Chicago, IL.

Cohen, J. (1977). Statistical power for the behavioral sciences. New York: Academic Press.

Cohen, J., & Cohen, P. (1975). Applied multiple regression/ correlation analysis for the behavioral sciences. New York: John Wiley & Sons.

Fanning, F., & Newman, I. (1974). The development and demonstration of multiple linear regression models for operant conditioning questions. Multiple Linear Regression Viewpoints, 4(4), 65-87.

Fraas, J., & Newman, I. (1978). Malpractice of statistical interpretation. Multiple Linear Regression Viewpoints, 9, 1-25.

Glass, G. V. (1976). Primary, secondary and meta-analysis of research. Educational Researcher, 5, 3-8.

Hedges, L. V., & Olkin, I. (1983). Regression models in research synthesis. The American Statistician, 37(2), 137-140.

Kelly, F. J., McNeil, K., & Newman, I. (1973). Suggested inferential statistical models for research in behavior modification. Journal of Experimental Education, 41(3).

Kraemer, H. C. (1983). Theory of estimation and testing of effect sizes: Use in meta-analysis. Journal of Educational Statistics, 8(2), 93-101.

McNeil, K., Kelly, F. J., & McNeil, J. (1975). Testing research hypotheses using multiple linear regression. Carbondale, IL: Southern Illinois University Press.

Newman, I., Williams, J., & Bobner, R. (1982, August). Dependent binary data: A Comparison of regression and chi-square solutions. Paper presented at the joint meetings of the American Statistical Association, Cincinnati, OH.

Pedhazur, E. J. (1982). Multiple Regression in Behavioral Research. New York: Holt, Rinehart & Winston.

Rosenthal, R. (1978). Combining results of independent studies. Psychological Bulletin, 85(1), 185-193.

Strube, M. J. (1985). Combining and comparing significance levels from nonindependent hypothesis tests. Psychological Bulletin, 97(2), 334-341.

Tracz, S. M., & Elmore, P. B. (1985, August). The issue of nonindependence in correlational meta-analyses. Paper presented at the joint meetings of the American Statistical Association, Las Vegas, NV.

Williams, J. (1983). A note on coding subjects effect in treatment X subject designs. In I. Newman & C. Benz Multiple linear regression: Readings, exams, problems and syllabus (pp. 97-102). Akron, OH: Multiple Linear Regression Special Interest Group of the American Educational Research Association.